EXHIBIT 11

2007 WL 5058726 (N.D.Cal.) (Expert Report and Affidavit)
United States District Court, N.D. California.

In Re BEXTRA AND CELEBREX MARKETING, Sales Practices and Products Liability Litigation.

No. M05-CV-01699.
March 30, 2007.

Expert Report of Professor Lee-Jen Wei

Name of Expert: Lee-Jen Wei, Ph.D.
Area of Expertise: Accounting & Economics >> Statistics
Representing: Defendant
Jurisdiction: N.D.Cal.

TABLE OF CONTENTS

I. INTRODUCTION

A. Qualifications

1. I received a Ph.D. degree in Statistics in 1975 from the University of Wisconsin. I have been a tenured professor of biostatistics at Harvard University since 1991 and a professor of biostatistical science and computational biology at Dana-Farber Cancer Institute, Harvard Medical School, since 1997. I have also served as the co-director of the bioinformatics core at Harvard School of Public Health since 2003. From 2003 to 2004, I served as acting chair of the department of biostatistics at Harvard University. I. was a tenured full professor of biostatistics and statistics at University of Wisconsin, University of Michigan, and George Washington University from 1982-1991. I was named Cancer Expert by The National Cancer Institute in 1980. Since 2004, I have been actively involved in cardiovascular disease clinical trial research with cardiologists at Brigham & Women's Hospital in Boston.

2. My scholarly writings include over 120 articles in peer-reviewed academic journals. I am responsible for developing numerous novel statistical methods for designing and analyzing clinical trial studies. Many of these methods have been included in the most commonly used statistical packages such as SAS and S-plus. I have served on the editorial boards of a number of statistical journals and am an elected Fellow of the American Statistical Association and Institute of Mathematical Statistics. I was named “Statistician of the Year” in 2007 by the Boston Chapter, American Statistical Association. I am a frequent reviewer for National Institutes of Health (“NIH”) grants and contracts and was on the Food and Drug Administration (“FDA”) advisory committee board.

3. I am an academic affiliate of Analysis Group, Inc. and I work with that organization on a regular basis on statistical analysis projects. My curriculum vitae, which includes a complete list of my publications is listed in Appendix A. My testimony in legal proceedings in the past four years is listed in Appendix B.

B. Assignment

4. I have been asked to determine whether Celebrex, at daily doses of 200 mg, 400 mg and 800 mg, is associated with a risk of specific cardiovascular (“CV”) events relative to placebo and non-selective non-steroidal anti-inflammatory drugs (“NSAIDs”) based on reliable data sets accessible to me from comparative clinical trials. [1]

5. All materials reviewed and used for the preparation of my report are listed in Appendix C.

II. SUMMARY OF OPINIONS

6. When the risk of CV events (see the definition below) for patients taking Celebrex is compared with those taking a placebo, based on my analysis, I find that:

i. For the 200 mg daily dose, there is no evidence of a statistically significant difference in CV risk between Celebrex and placebo.

Back to top

ii. For the 400 mg daily dose, there is no evidence of a statistically significant difference in CV risk between Celebrex and placebo.

iii. For the 800 mg daily dose, there is some evidence of a statistically significant difference in CV risk between Celebrex and placebo.

7. When the risk of CV events for patients taking Celebrex is compared against those taking other non-selective NSAIDs, based on my analysis, I find that

i. For the 200 mg daily dose, there is no evidence of a statistically significant difference in CV risk between Celebrex and non-selective NSAIDs.

ii. For the 400 mg daily dose, there is no evidence of a statistically significant difference in CV risk between Celebrex and non-selective NSAIDs.

iii. For the 800 mg daily dose, there is no evidence of a statistically significant difference in CV risk between Celebrex and non-selective NSAIDs.

III. PRIMER ON STATISTICAL ANALYSIS, CLINICAL TRIALS, AND META-ANALYSIS

A. Making Inferences About Population Characteristics Using a Sample of Subjects

8. Suppose that we are interested in the incidence rate of a certain clinical event (for example, CV events) among patients treated with Celebrex relative to the corresponding rate for patients who have not been exposed to the drug. To do this, we take a sample from a population of patients treated with Celebrex and another sample from the population of patients who did not receive Celebrex. Assuming that these samples are valid representatives of the two populations, statistical methods can be used to determine whether the Celebrex group has a different rate of CV events than the non-Celebrex group. Since we draw conclusions based on a subsets of patients only, any qualitative or quantitative interpretation of the result (i.e., whether the rate is higher or not) is subject to so-called sampling error. In other words, the observed incidence rate may be higher (leading to a possible false positive finding) or lower (leading to a possible false negative finding) than the true incidence rate in the population. An efficient statistical method for analyzing such data minimizes the chance of making these two types of errors. It is important to note that except for the usage of Celebrex, ideally the Celebrex users in the sample should be similar to the patients in the non-Celebrex sample with respect to important observable or unobservable confounders (e.g., age, disease status).

9. Once an investigator has determined the patient population of interest, she/he must draw a valid sample from the population. Next, the investigator has to determine what clinical or biomarker endpoints are most appropriate to quantify the effect of the treatment (with respect to efficacy or toxicity). Let us assume that the investigator uses certain types of CV events as endpoints to evaluate potential Celebrex toxicity. Suppose that based on a sample of 100 patients, four patients experience CV events, An obvious estimate of the incidence rate of toxicity for the underlying population is 0.04 (or 4%). This is called a point estimate. However, this estimate is based on a relatively small set of patients. The true incidence rate for the entire population may be more or less than 0.04. Another investigator using a different sample or study may find that none of the patients experienced a CV event. Therefore, when observing results from a single sample, it is important to attach a level of confidence to the observed point estimate. This quantitative, scientific process is called “drawing inferences” about the true incidence rate.

10. It is important to report not only the point estimate but also the confidence interval around the point estimate as a measure of the precision of the single point estimate. The wider the confidence interval is, the larger the range of possible values for the underlying true population effect around the point estimate will be.

11. A commonly used method for inference is to define a 95% confidence interval for the true incidence rate. Based on the above sample (four patients had CV events among 100 patients), the 95% confidence interval for the true rate includes all possible values between the lower bound of 0.001 and the upper bound of 0.079 and is denoted by (0.001, 0.079). That means that if we were to repeat the analysis using different independent samples from the same patient population, say, 1000 times and build 95% confidence intervals around the point estimate each time, 950 the resulting 1000 confidence intervals would include the true incidence rate (950 out of 1000 is 95%). Since we only

have a single dataset and a single confidence interval, we may consider that the observed interval includes the truth with 95% chance. Or, loosely speaking, we can say that with 95% chance, the true incidence rate is between 0.001and 0.079.

12. Now, suppose that we are interested in comparing the incidence rates across two groups of patients with respect to certain CV events. Suppose that the first group of patients is treated with Celebrex (as discussed previously), and the second group of patients does not take any active drug (placebo). Suppose that in a sample of 100 placebo patients we observe two CV events. The point estimate of the CV events incidence rate in the underlying population not taking Celebrex is 0.02 (which is smaller than the hypothetical observed incidence rate of 0.04 for the Celebrex group). Can we conclude that patients receiving Celebrex have an increased likelihood of experiencing a CV event relative to patients who are not taking the drug? The answer is no. In fact, the 95% confidence interval for the difference between two true incidence rates of 0.02 (0.04 for the Celebrex arm minus 0.02 for the placebo arm) ranges from negative 0.03 to positive 0.07. That is, based on this data, we cannot say that Celebrex is associated with an elevated risk of CV events relative to placebo.

13. The confidence interval estimation procedure that I have just described may be used to test hypotheses about the relative CV incidence rate among patients administered Celebrex relative to patients administered a placebo. If the 95% confidence interval for the difference between the two rates includes the null hypothesis ("null") value zero, we cannot conclude that Celebrex is associated with a statistically significant difference with respect to this endpoint relative to placebo. On the other hand, if the confidence interval does not include zero, a true statistical difference between the two groups is likely. Naturally, since we only have limited data, this claim is subject to error. Here, the probability of making a falsely positive statement (that is, of stating that there is a difference between the Celebrex and the placebo group even though there is none) is 5% (the conventional acceptable error probability in the literature for a single endpoint analysis). Statisticians then minimize the probability of a false negative finding (that is, of stating that there is no difference between the two arms even though one exists). It is important to note that even if we find that there is a statistically significant difference between the two arms, the next important question is whether that difference is clinically meaningful with respect to risk-benefit considerations.

14. The 95% level for the confidence interval, or the 5% level of significance for testing hypothesis, is typically used by investigators and statisticians to establish the "statistical significance" of a result when testing a single clinical primary endpoint. This level can be quite "liberal" (i.e., can result in statements of statistical significance when none exists) if multiple endpoints are examined simultaneously. For example, in the clinical trial investigating the use of Celebrex in colon cancer prevention, the primary endpoint was the efficacy of the drug (with respect to cancer). However, numerous potential adverse event endpoints were also examined. Using the 5% rule for claiming statistical significance to analyze simultaneously a large number of endpoints in a study will yield a high rate of false positive findings across all endpoints. It is not unusual that when the 5% test hypothesis rule is applied simultaneously to a number of endpoints in a study, the overall false positive rate is as high as 20% (that is, a very high chance of claiming a drug is not safe with respect to the placebo, when, in fact, there is no difference between two groups).

B. Using Clinical Trials to Evaluate a Drug Effect (Efficacy or Toxicity)

15. Let me turn to the issues of comparing two groups of patients, one receiving Celebrex and the other receiving a placebo. To make sure that two samples of patients (e.g., Celebrex and placebo) are comparable with respect to all potential confounders, investigators often rely on a randomized clinical trial setting. Such a medical study yields a well-designed experiment for generating reliable prospective data on drug efficacy or safety. Such studies are conducted and monitored according to a pre-specified protocol which details the treatments administered (e.g., form, dosage, frequency), the clinical or biological endpoints (e.g., lab value, patient's quality of life, time to remission, time to a toxicity event), the study patient population (e.g., elderly, suffering from rheumatoid arthritis) and other clinical and statistical considerations. The trial is usually randomized and blind. Patients are assigned randomly to one of the study arms and neither physicians nor patients are told whether the patient is receiving an active drug (Celebrex) or a placebo. This avoids selection bias or other experimental bias. Results from a well-conducted, randomized clinical trial are regarded as a gold standard in controlled settings to evaluate the efficacy and safety of a treatment.

16. Investigators may also be interested in comparing several groups of patients simultaneously who are receiving different doses of the active treatment. Furthermore, the investigators may wish to compare the active treatment to an alternative (perhaps “older”) treatment rather than a placebo.

17. Suppose that the primary endpoint of a study is the efficacy of Celebrex relative to placebo. At the end of the study (or at the interim analysis), a summary statistic is constructed to estimate the relative efficacy of Celebrex and placebo. In other words, the difference between the two groups is quantified. Then, a 95% confidence interval is calculated around this difference. If the confidence interval excludes the “null” value (usually it is zero), the two groups are statistically significantly different from each other with respect to the primary efficacy endpoint. Naturally there are often numerous other secondary safety endpoints considered in the study. However, when “the statistical significance definition” (for example, 5% rule for testing hypothesis) is applied to other secondary endpoints, we usually interpret the results with great caution. That is, if for a specific endpoint, we find that the result is “statistically significant”, we usually use a much higher confidence level, say, 99% instead of 95% to avoid the so-called multiple comparison problem.

18. A drug may also be studied under various settings including different patient populations, dosing, or study duration. Moreover, the primary clinical endpoints may vary from study to study. It is not uncommon for different studies to yield different and even contradictory results, especially with respect to secondary endpoints. Therefore, drawing conclusions from a single study with respect to a non-primary endpoint is speculative and should be interpreted with great caution.

C. Examining the Safety of a Drug - Identification Issues with Small Numbers of Events

19. While clinical trials are often geared towards assessing the efficacy of a particular product, other relevant outcomes are also routinely collected by investigators. Of particular interest is the collection of adverse events and whether or not they are associated with the drug being studied. In fact, the words “adverse events” have a broad meaning. The occurrences of these clinical events (for example, CV events) may not be unusual at all for the general study population. Moreover, such “adverse events” may be related to a host of factors including disease progression and have nothing to do with the study drug.

20. An investigator's ability to detect the difference in adverse events among study arms depends largely on the magnitude and frequency of adverse events. At one end of the spectrum are drug-associated events that are easy to detect. Some drugs have very high rates of associated adverse events resulting in an obvious difference relative to placebo. However, even a low rate of adverse events among patients receiving the active treatment may be easy to detect if the adverse event is virtually never observed in the general population and is clearly identifiable as a “signature” event for the drug at issue. At the opposite end of the spectrum, if a particular adverse event is common in the population and is observed to occur at very similar rates in the two groups, detecting whether any difference observed is a true difference that exists in the underlying population or is simply the result of chance will require large samples of patients. Similarly, if a particular event is rarely observed and is not exclusively associated with the drug at issue, determining whether any elevated adverse event rates among patients receiving the drug truly exist in the underlying population will require large sample sizes.

21. In addition, if the primary endpoint of the study is not safety, conclusions about drug safety based on a single study may be misleading. For example, the follow-up times between two groups of patients may be different because of differential efficacy across the two groups (patients receiving the placebo may quit early to receive other treatments). This may cause imbalanced follow-up patterns for safety data across groups. Meta-analysis techniques (described below) provide a rigorous approach to combine drug-specific data across multiple studies. A consistent pattern of safety or efficacy findings over numerous studies once combined using scientifically sound meta-analysis methods will help determine whether usage of a drug results in elevated risks.

D. Quantifying the Difference in Adverse Event Rates across Two Patient Groups

22. Using randomized controlled trials, investigators may measure any differences in adverse events using a variety of metrics. For example, if the investigator observes a 0.01 rate of adverse events in the treatment arm and a 0.008 rate

of adverse events in the placebo arm, the difference (called "risk difference," denoted by RD) would be 0.01-0.008 = 0.002. A risk difference of 0 indicates that there is no difference in the likelihood of adverse events between patients in the two arms of the study. As explained earlier, a risk difference greater than zero in the sample studied may or may not indicate that the drug is associated with an elevated risk in the underlying population depending on the 95% confidence interval around the point estimate of the risk difference. If the 95% confidence interval includes 0, then the finding is not statistically significant and the associated p-value is greater than 0.05. If the 95% confidence interval excludes zero, then the finding is statistically significant and the p-value is less than 0.05.

23. Alternatively, the investigator could report the ratio of the two rates. This is called the "risk ratio" ("RR") and, in this example, it would be 0.01 / 0.008 = 1.25. Another commonly used metric in the literature is the "odds ratio." In my example, patients in the treatment arm have a 1 to 99 odds of experiencing an adverse event (they have 1% chance of experiencing the adverse event and 99% chance not to) while patients in the placebo arm have a 1 to 124 odds of experiencing an adverse event. The odds ratio = (1/99)/(1/124) =1.253. In this example, the odds ratio (1.253) and the risk ratio (1.25) are very close because the likelihood of events is small. An odds ratio or a risk ratio of 1 indicates that there is no difference in the likelihood of an adverse event between patients in the two arms of the study. Similar to the risk difference, an estimated risk ratio that is different from one, based on the study results, may or may not indicate that the drug is associated with an elevated risk in the underlying population depending on whether the confidence interval associated with the ratio includes one or not (or the p-value associated with the difference is less than 0.05). In the rest of my analysis, I use both the risk difference and the risk ratio as measures of the contrast in adverse events observed across study arms.

E. Meta-Analysis: Looking at the Totality of Evidence

24. When very small numbers of adverse events are observed, conclusions based on a single study are likely to be unreliable, especially for non-primary safety endpoints. By chance alone, one may find an apparent statistically significant difference in adverse event rates in a single study even though no difference truly exists in the underlying population. The opposite may also be true when a single study is examined. Therefore, conclusions should be drawn based on the totality of evidence available from multiple studies providing consistent results. Combining such studies using meta-analysis techniques is a valid scientific way to consider the totality of evidence, review the consistency of results and draw appropriate conclusions. Statistically significant findings repeatedly observed across multiple independent studies and confirmed using meta-analysis techniques provide investigators confidence that the finding is real, rather than a statistical artifact. On the other hand, if a statistically significant finding from a single study is not repeatable across multiple independent studies and is not confirmed by a meta-analysis, the investigator may conclude that the effect may be a statistical artifact that does not indicate a true effect in the underlying population.

25. By combining studies, a meta-analysis increases the overall sample considered and may increase the number of events observed thereby increasing the power of detecting potential safety signals. Statistically, a meta-analysis combines multiple studies, each contributing a point estimate for the difference in effect across arms and a confidence interval around this difference. Combining these confidence intervals across studies provides a global assessment of the treatment difference with a tighter (smaller) confidence interval and, therefore, a greater certainty in the statistical inference about the true difference. In other words, meta-analyses combine data from multiple trials to identify important conclusions that may not be observable based on any single trial.

26. Combining studies in a meta-analysis can be done using either "fixed effects" or "random effects" modeling approaches. [2] The fixed effects model assumes that the true contrast across study arms (for example, the risk ratio or risk difference) is constant across all individual studies even though the estimated contrasts may vary across studies due to sampling variation. Alternatively, if the investigator believes that the true contrasts may vary across studies, a random effects model is more appropriate to combine data across the studies. Since the random effects model does not require the assumption that true contrasts be the same, meta-analysis results based on the random effects model are usually more reliable. On the other hand, the random effects model assumes a parametric distribution for the random effects and is, therefore, still subject to modeling assumptions. Therefore, my practice is to do both

Back to top

analyses; if the results are quite different, further model checking should be performed to examine which approach is more appropriate.

27. There are additional methodological issues to consider when performing a meta-analysis that includes studies with no events in at least one, and in some cases both arms. An odds ratio cannot be calculated unless events are observed in both arms. Similarly, a risk ratio cannot be calculated if there are no events in the arm of the study that provides the denominator of the ratio. This is a significant issue in this litigation since the CV events of interest are rare and non existent in most clinical trials comparing Celebrex and placebo (or NSAIDs). Ignoring studies with no CV events would yield misleading conclusions on the safety of Celebrex. For example, suppose that one observes 10 studies comparing drug A and drug B; none of the 10 studies shows adverse events for patients receiving drug A, but they do indicate adverse events for patients receiving drug B. Even though the odds or risk ratios cannot be directly calculated and analyzed in this case, it is quite obvious that these studies contain potentially useful information about the relative safety of the two drugs. Similarly, suppose that 9 out of the 10 studies show no events in either arm, while the 10th study shows one event for drug A and two events for drug B. Limiting the analysis to just the 10th study may overstate the true difference between the two arms. The fact that the other 9 studies showed no adverse effect with either drug is relevant information that should be incorporated in the analysis. I would like to emphasize that information about the underlying adverse effect rates for both arms is very important for the decision making process.

28. To include studies which do not have events in one or both arms, researchers commonly add a small positive number (for example, 0.5) to all outcome counts in a study. For example, suppose that 1 of 50 patients taking drug A experienced a CV event, while none of 50 patients taking drug B experienced any. Then, a value of 0.5 would be added to each group of patients. [3] The risk ratio can now be calculated as (1.5/51)/(0.5/51)=3.

29. This “fixed imputation” method has a drawback in the case of unbalanced study arms (where the numbers of patients are not the same in both arms). For example, suppose that no adverse events were observed in a trial including 50 patients taking drug A and 100 patients taking drug B. Adding a fixed amount of 0.5 will lead to a risk ratio of approximately 2 even though the risk of events (zero) is equal in both arms and should yield a risk ratio of one. One way to avoid this problem is to add values proportional to the number of subjects in each arm. In the present example, one would add 50/(50*100)=0.33 events to the patient groups receiving drug A and 100/(50*100)=0.67 events to the patient groups receiving drug B. Then the estimated risk ratio would be (0.33/50.33)/(0.67/101.67)=1.

30. The risk difference (the difference between event rates across two arms) offers a better summary of the contrast than the risk ratio particularly when many studies have no events. Because this measure is a simple difference between the two event rates, it is well-defined even when both study arms have zero events: in such a case the risk difference will be zero (zero minus zero). Therefore, the calculation of risk difference requires no imputation. However, when constructing the confidence interval around this estimate, some imputation may still be required because standard methods to calculate confidence intervals cannot be used. I explain how to perform these imputations in the results section.

31. It is also possible to avoid imputations altogether when calculating confidence intervals around risk differences by constructing “exact confidence intervals.” [4] These exact confidence intervals can then be combined in a meta-analysis using the technique proposed recently by Singh, Xie and Strawderman (2005) [5] . I am currently directing the development of a computer code to implement this method to corroborate my current findings which are based on the imputation method.

IV. Implementing a Meta-Analysis to Analyze CV Events Using Randomized Clinical Trials

32. This section of my report describes the meta-analysis implemented to examine (1) whether patients treated with Celebrex and patients treated with placebo have the same risk of experiencing a CV event; and (2) whether patients treated with Celebrex and patients treated with non-selective NSAIDs have the same risk of experiencing a CV event. Following Dr. Milton Packer's instructions, I focus on a composite clinical CV event endpoint (“APTC”), which was defined by the Antiplatelet Trialist's Collaboration. APTC includes non-fatal myocardial infarctions (“MI'), non-fatal

Back to top

strokes ("Strokes"), and CV deaths ("Deaths"). I also report results for MIs, Strokes and Deaths separately. Following Dr. Milton Packer's instructions, I have analyzed data separately for patients receiving Celebrex daily doses of 200 mg, 400 mg, and 800 mg.

33. The studies included in the meta-analysis are summarized in the next section. I report all results both in terms of risk ratio and risk difference to provide a complete summary of the contrast (or difference) in CV events between patients receiving Celebrex and patients receiving a placebo or non-selective NSAIDs.

A. Selection of Studies for Inclusion in the Meta-Analysis

34. Exhibit 1 of my report describes the study selection process. I looked into three main types of studies: Pfizer-sponsored clinical studies, Investigator Initiated Research ("IIR"), and government-sponsored studies. In addition, I am also performing an ongoing literature review.

35. For the Pfizer-sponsored studies, my initial source of data is the Right Track database of clinical trial protocols and studies. [6] When Pfizer acquired Pharmacia in April 2003, all the clinical trial protocols and studies regarding Celebrex contained in Pharmacia's Starbase were incorporated into Pfizer's Right Track. At the time of my analysis I was provided with a list of 325 protocols and studies involving Celebrex. [7]

36. From these 325 protocols and studies, the staff members at Analysis Group, under my supervision, selected 232 that were flagged as "approved", "terminated" or "completed." [8] These flags correspond to protocols that have been approved, studies that have been terminated, and studies that have been completed.

37. Of the 232 studies remaining 106 had reports available and in English, which was necessary for gathering the relevant information for my investigation. [9] From these 106 studies I selected those that were double-blind, randomized, placebo or non-selective NSAID controlled, that reported myocardial infarction events, stroke events, cardiovascular deaths or APTC, number of patients by dose, and treatment duration of more than one week. [10] Of the 106 studies, 83 met my criteria.

38. I also considered the trials from Pfizer's IIRs. These protocols and study reports are stored in two databases at Pfizer. In the U.S. before January 2006, the Pfizer data were stored in the Non-Registry database which is part of the Right Track database. More recent IIRs as well as legacy IIRs (e.g., Pharmacia, Warner-Lambert) are stored in the Independent Grant database ("IG"). I received from Pfizer a list of 361 protocol and study names corresponding to all Celebrex studies in the Non-Registry database [11] and a list of 158 protocol and study names corresponding to all Celebrex studies in the IG database. [12] The titles of Celebrex protocols and studies listed in the IG and Non-Registry database were not informative enough for me to apply my selection criteria. In addition, I received 46 IIR abstracts, publications and study reports. [13] I reviewed these 46 abstracts, publications and study reports and was able to identify 8 IIRs which met the criteria of being double-blind, randomized, placebo or NSAID controlled, having reported cardiovascular events such as myocardial infarction, stroke, cardiovascular death and APTC, number of patients by dose, duration of more than one week, and being available and in English. This research is ongoing. I will continue to update my analyses as new protocols and studies are produced.

39. Finally, I considered the government-sponsored studies. I searched the NIH website [14] for all Celebrex studies and identified 170 studies. [15] I reviewed the project descriptions and publications where available and relevant. I used similar criteria as in the selection process of the IIRs, From the 170 studies I identified three studies which met my criteria, ADAPT, APC, and PreSAP.

40. The IIR and NIH additions bring the total number of trials to 94 (83 * 8 * 3 = 94). Of the 94 selected trials, I further focused my selection by including only studies that lasted at least two weeks, for a final selection of 73 studies.

41. Dr. Milton Packer examined these 73 studies and provided me with a count of myocardial infarctions, strokes, and cardiovascular deaths that were mutually exclusive. For the ADAPT, APC, and PreSAP studies, which have already been published and previously adjudicated, Dr. Packer decided to use their reported APTC events number directly because from the publications he could not determine with certainty whether the myocardial infarctions, strokes

cardiovascular deaths had been reported consistently as mutually exclusive events. From these 73 reports, Dr. Packer provided me with the number of adjudicated CV events and patients for 64 reports. [16] , [17]

42. There are a number of meta-analyses that have examined the difference in CV events between Celebrex and other comparators. Most of these analyses utilized study populations that largely overlapped with each other and that are included in my analysis. To the best of my knowledge, the set of studies which I am considering is the most complete up-to-date set analyzed in the context of a meta-analysis of Celebrex trials.

43. Because Celebrex is currently on the market, additional studies and data are likely to become available in the future, I intend to search the public literature and to obtain all new study data and other non-Pfizer data that can be identified. As these data accumulate, I will update my results accordingly.

B. Meta-Analysis

44. Before numerically combining individual study results, first, I examine the risk difference for each study, by dose (200 mg, 400 mg, and 800 mg). I constructed a 95% confidence interval for each study and provided a graphical display to examine whether there is a visually consistent pattern across all the studies with respect to the CV event rate (APTC in this case). This type of graphical summary is quite informative. After the graphical representation, for risk difference and risk ratio, I then numerically combine individual study results using standard meta-analysis methods. As I have outlined in section III, one issue to consider is the rarity of observed CV events. For example, out of the 24 studies that are available to analyze the CV risk associated with 400 mg doses of Celebrex relative to placebo, only 5 studies have APTC events in both arms. The vast majority (79%) of studies have zero events in at least one arm ("zero-arm studies"). One potential approach is to ignore all the zero-arm studies and focus only on these 5 studies. [18] However, if we ignore the zero-arm studies in the analysis, the results may be misleading because it ignores the information contained in the zero-arm studies. Following standard practices in meta-analysis, I utilize the proportional imputation method described above for the zero-arm studies as my primary analysis tool.

45. Additionally, as I have discussed above, there are a number of methodological approaches that one might undertake for performing a meta-analysis. For my primary analysis, I use the Mantel-Haenszel ("MH") weighting scheme for the risk ratio and risk difference. [19] Moreover, I used the random effects modeling approach to avoid the a priori assumption of the fixed effects model. [20]

46. I also performed sensitivity analysis using an MH fixed effects model, fixed and random effects inverse variance ("IV") models, and alternative imputation methods (no imputation and fixed imputation). Results from these sensitivity analyses are presented in Appendix D of my report. I find that all the methods conventionally used for meta-analysis give very similar results to those based on the MH random effects approach for risk ratio and risk difference. To summarize results, for each risk measure (risk ratio or difference), I report the point estimates and the corresponding combined 95% confidence intervals.

V. RESULTS

47. In this section I present results from my analyses. The following Sub-section A presents results based on studies that compared Celebrex with a placebo, including four sections describing results for alternative measures of CV events: APTC (sub-section A.i), Strokes (sub-section A.ii), MIs (sub-section A.iii), and Deaths (sub-section A.iv). Sub-section B presents my results from comparing Celebrex to non-selective NSAIDs for the same four measures of CV events; APTC (sub-section B.i), Strokes (sub-section B.ii), MIs (sub-section B.iii), and Deaths (sub-section B.iv).

A. Comparing CV Events among Patients Receiving Celebrex or Placebo

48. Exhibit 2 presents data on study sample sizes and administered doses. [21] There are 30 studies including a comparison of placebo to 200 mg administration of Celebrex, 24 for a comparison to 400 mg dose, and 9 for a comparison to 800 mg dose. The 200 mg studies are all less than 16 weeks in duration. The 400 mg studies include three studies - APC, PreSAP, and ADAPT - which have durations of 156 weeks. The 800 mg studies also have the maximum duration of 156 weeks.

Back to top

i. Analysis of APTC Events

49. First, for the risk difference, for each dose (200 mg, 400 mg, and 800 mg), I constructed the exact 95% confidence interval for each study and provided a graphical display to identify visually a possible consistent pattern across all the studies with respect to the CV event rate. These confidence intervals are presented using the "tree diagrams" in Exhibits 3A through 3C. In Exhibit 3A, each horizontal line denotes the exact 95% confidence interval for the risk difference with respect to APTC (Celebrex minus placebo) from a single study for patients receiving Celebrex 200 mg/daily. All the studies, even those without CV events, are included in this display. It is clear that all the confidence intervals contain zero suggesting that there is no difference in the rate of APTC across the Celebrex and placebo groups. In Exhibit 3B, I display a similar plot for all studies including patients receiving Celebrex 400 mg/daily. All the exact confidence intervals contain zero except for one study (APC). In Exhibit 3C, I display a similar plot for all studies including patients receiving Celebrex 800 mg/daily. All the intervals include zero except for the APC study.

50. Exhibit 4 presents the results for APTC based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.88 (Celebrex over placebo). The corresponding 95% confidence interval is (0.46, 1.72) which contains the null value one. For risk difference (RD), the point estimate is - 0.0002 and its 95% confidence interval is (-0.0025, 0.0022), which contains the null value zero. Thus, the point estimate suggests a lower APTC risk associated with Celebrex. However, that difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of APTC events between patients receiving Celebrex 200 mg/daily and patients receiving placebo. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 1.32 with a 95% confidence interval of (0.93, 1.86), which contains the null value one. For RD, the point estimate is 0.0007 with a 95% confidence interval of (-0.0022, 0.0036), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of APTC events between patients receiving Celebrex 400 mg/daily and patients receiving placebo. For studies including patients receiving Celebrex 800 mg/daily, the point estimate of RR is 2.30 with a 95% confidence interval of (1.13, 4.70). For RD, the point estimate is 0.0047 with a 95% confidence interval of (-0.0034, 0.0128). Since the RR is statistically significantly different from one (the 95% confidence interval does not include one) but the RD is not statistically significant, I conclude that there is some evidence that patients receiving Celebrex at a dose of 800 mg/daily may experience a higher rate of APTC events than patients receiving placebo.

51. In addition to my main analysis, I also performed a number of sensitivity analyses using the APTC measure. First, I only conducted this analysis on studies which do not require imputation. [22] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 1 (Risk Ratio) and Appendix D Exhibit 2 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. There is no statistically significant difference in the rate of CV events for patients receiving Celebrex 200 mg/daily or 400 mg/daily versus patients receiving placebo. The sensitivities also indicate that there is mixed evidence that patients receiving Celebrex 800 mg/daily may experience a higher rate of APTC events than patients receiving placebo.

ii. *Analysis of Strokes*

52. Exhibit 5 presents the results for Strokes based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.93 (Celebrex over placebo). The corresponding 95% confidence interval is (0.46 1.88) which contains the null value one. For risk difference (RD), the point estimate is - 0.0001 and its 95% confidence interval is (-0.0022, 0.0020), which contains the null value zero. Although the point estimate suggests a lower risk of Strokes associated with Celebrex, the difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of Strokes between patients receiving Celebrex 200 mg/daily and patients receiving placebo. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 1.06 with a 95% confidence interval of (0.60, 1.88), which contains the null value one. For RD, the point estimate is -0.0004 with a 95% confidence interval of (-0.0029, 0.0021), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of Strokes between patients receiving Celebrex 400 mg/daily and patients receiving placebo. For the 800 mg/daily, the point

estimate of RR is 1.37 with a 95% confidence interval of (0.51, 3.67). For RD, the point estimate is 0.0012 with a 95% confidence interval of (-0.0033, 0.0057). Again, I do not find any statistically significant difference in the rate of Strokes between patients receiving Celebrex 800 mg/daily and patients receiving placebo.

53. In addition to my main analysis, I also performed a number of sensitivity analyses using the Strokes measure. First, I only conducted this analysis on studies which do not require imputation.[23] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 3 (Risk Ratio) and Appendix D Exhibit 4 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. There is no statistically significant difference in the rate of Strokes for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving placebo.

iii. ***Analysis of MIs***

54. Exhibit 6 presents the results for MIs based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.97 (Celebrex over placebo). The corresponding 95% confidence interval is (0.48 1.96) which contains the null value one. For risk difference (RD), the point estimate is 0.0000 and its 95% confidence interval is (-0.0022, 0.0021), which contains the null value zero. Thus, the 95% confidence interval suggests that the difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of MIs between patients receiving Celebrex 200 mg/daily and patients receiving placebo. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 1.60 with a 95% confidence interval of (0.89, 2.91), which contains the null value one. For RD, the point estimate is 0.0014 with a 95% confidence interval of (-0.0011, 0.0038), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of MIs events between patients receiving Celebrex 400 mg/daily and patients receiving placebo. For the 800 mg/daily, the point estimate of RR is 1.75 with a 95% confidence interval of (0.70, 4.41). For RD, the point estimate is 0.0026 with a 95% confidence interval of (-0.0028, 0.0081). Again, I do not find any statistically significant difference in the rate of MIs between patients receiving Celebrex 800 mg/daily and patients receiving placebo.

55. In addition to my main analysis, I also performed a number of sensitivity analyses using the MIs measure. First, I only conducted this analysis on studies which do not require imputation.[24] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 5 (Risk Ratio) and Appendix D Exhibit 6 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis broadly confirms my conclusions. For the fixed and proportional imputation methods, there is no statistically significant difference in the rate of MIs for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving placebo across risk measures. For the method with no imputation involved, which excludes the majority of the studies, there is no statistically significant difference in the rate of MIs for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving placebo using risk ratio. For risk difference, there is no statistically significant difference in the rate of MIs for patients receiving Celebrex 200 mg/daily or 800 mg/daily versus patients receiving placebo, but for 400 mg/daily there is a statistically significant difference in the rate of MIs for patients receiving Celebrex versus patients receiving placebo. This isolated case, which excludes most of the studies, does not alter my conclusion based on my aforementioned primary analysis regarding the safety issues with 400 mg dose.

iv. ***Analysis of CV Death***

56. Exhibit 7 presents the results for Deaths based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.95 (Celebrex over placebo). The corresponding 95% confidence interval is (0.47 1.95) which contains the null value one. For risk

difference (RD), the point estimate is 0.0001 and its 95% confidence interval is (-0.0019, 0.0021), which contains the null value zero. Thus, the 95% confidence interval suggests that the difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of Deaths between patients receiving Celebrex 200 mg/daily and patients receiving placebo. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 1.14 with a 95% confidence interval of (0.65, 2.02), which contains the null value one. For RD, the point estimate is 0.0006 with a 95% confidence interval of (-0.0015, 0.0027), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of Deaths between patients receiving Celebrex 400 mg/daily and patients receiving placebo. For the 800 mg/daily, the point estimate of RR is 1.72 with a 95% confidence interval of (0.59, 4.99). For RD, the point estimate is 0.0027 with a 95% confidence interval of (-0.0018, 0.0072). Again, I do not find any statistically significant difference in the rate of Deaths between patients receiving Celebrex 800 mg/daily and patients receiving placebo.

57. In addition to my main analysis, I also performed a number of sensitivity analyses using the Deaths measure. First, I only conducted this analysis on studies which do not require imputation. [25] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 7 (Risk Ratio) and Appendix D Exhibit 8 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. There is no statistically significant difference in the rate of Deaths for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving placebo.

B. Comparing CV Events among Patients Receiving Celebrex or non-selective NSAIDs

58. Exhibit 8 presents data on study sample sizes and administered doses. [26] There are 26 study arms for the 200 mg, 19 for the 400 mg, and 3 for the 800 mg dose.

i. *Analysis of APTC Events*

59. First, for the risk difference, for each dose (200 mg, 400 mg, and 800 mg), I construct the exact 95% confidence interval for each study and provide a graphical display to identify visually a possible consistent pattern across all the studies with respect to the CV event rate. These confidence intervals are presented using the "tree diagrams" in Exhibits 9A through 9C. In Exhibit 9A, each horizontal line denotes the exact 95% confidence interval for the risk difference with respect to APTC (Celebrex minus non-selective NSAIDs) from a single study for patients receiving Celebrex 200 mg/daily. All the studies, even those without CV events, are included in this display. All the confidence intervals contain zero suggesting that there is no difference in the rate of APTC across the Celebrex and non-selective NSAIDs groups. In Exhibit 9B, I display a similar plot for all studies including patients receiving Celebrex 400 mg/daily. All the exact confidence intervals contain zero. In Exhibit 9C, I display a similar plot for all studies including patients receiving Celebrex 800 mg/daily. All the intervals are also centered around zero.

60. Exhibit 10 presents the results for APTC based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.83 (Celebrex over non-selective NSAIDs). The corresponding 95% confidence interval is (0.52, 1.31) which contains the null value one. For risk difference (RD), the point estimate is -0.0003 and its 95% confidence interval is (-0.0018, 0.0011), which contains the null value zero. Thus, the point estimate suggests a lower APTC risk associated with Celebrex but that difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of APTC events between patients receiving Celebrex 200 mg/daily and patients receiving non-selective NSAIDs. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 0.83 with a 95% confidence interval of (0.54, 1.28), which contains the null value one. For RD, the point estimate is -0.0002 with a 95% confidence interval of (-0.0017, 0.0014), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of APTC events between patients receiving Celebrex 400 mg/daily and patients receiving non-selective NSAIDs. For the 800 mg/daily, the point estimate of RR is 0.88 with a 95% confidence interval of (0.54, 1.44) which contains the null value one. For RD, the point estimate is -0.0008 with a 95% confidence interval of (-0.0042,

0.0026), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of APTC events between patients receiving Celebrex 800 mg/daily and patients receiving non-selective NSAIDs.

61. In addition to my main analysis, I also performed a number of sensitivity analyses using the APTC measure. First, I only conducted this analysis on studies which do not require imputation. [27] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 9 (Risk Ratio) and Appendix D Exhibit 10 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. There is no statistically significant difference in the rate of CV events for patients receiving Celebrex 200 mg/daily, 400 mg/daily, or 800 mg/daily versus patients receiving non-selective NSAIDs.

ii. ***Analysis of Strokes***

62. Exhibit 11 presents the results for Strokes based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.67 (Celebrex over non-selective NSAIDs). The corresponding 95% confidence interval is (0.34, 1.29) which contains the null value one. For risk difference (RD), the point estimate is -0.0006 and its 95% confidence interval is (-0.0015, 0.0002), which contains the null value zero. Thus, the point estimate suggests a lower risk of Strokes associated with Celebrex but that difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of Strokes between patients receiving Celebrex 200 mg/daily and patients receiving non-selective NSAIDs. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 1.29 with a 95% confidence interval of (0.62, 2.69), which contains the null value one. For RD, the point estimate is 0.0006 with a 95% confidence interval of (-0.0006, 0.0018), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of Strokes events between patients receiving Celebrex 400 mg/daily and patients receiving non-selective NSAIDs. For the 800 mg/daily, the point estimate of RR is 0.36 with a 95% confidence interval of (0.13, 1.01). For RD, the point estimate is -0.0020 with a 95% confidence interval of (-0.0038, -0.0001) which does not contain the value of zero and is negative suggesting that Celebrex 800mg/daily users have a lower risk of Strokes. In general, I do not find any statistically significant difference in the rate of Strokes between patients receiving Celebrex 800 mg/daily and patients receiving non-selective NSAIDs.

63. In addition to my main analysis, I also performed a number of sensitivity analyses using the Strokes measure. First, I only conducted this analysis on studies which do not require imputation. [28] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 11 (Risk Ratio) and Appendix D Exhibit 12 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. For some cases, I find a slightly lower rate of Strokes for patients taking Celebrex but overall, there is no statistically significant difference in the rate of Strokes for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving non-selective NSAIDs.

iii. ***Analysis of MIs***

64. Exhibit 12 presents the results for MIs based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 1.09 (Celebrex over non-selective NSAIDs). The corresponding 95% confidence interval is (0.56, 2.13) which contains the null value one. For risk difference (RD), the point estimate is 0.0009 and its 95% confidence interval is (-0.0002, 0.0019), which contains the null value zero. Thus, the 95% confidence interval suggests that the difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of MIs between patients receiving Celebrex 200 mg/daily and patients receiving non-selective NSAIDs. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 1.15 with a 95% confidence interval of (0.52, 2.57), which contains the null value one. For RD, the point estimate is 0.0002 with a 95% confidence interval of (-0.0005, 0.0009), which

contains the null value zero. Again, I do not find any statistically significant difference in the rate of MIs between patients receiving Celebrex 400 mg/daily and patients receiving non-selective NSAIDs. For the 800 mg/daily, the point estimate of RR is 1.49 with a 95% confidence interval of (0.76, 2.94). For RD, the point estimate is 0.0015 with a 95% confidence interval of (-0.0011, 0.0041). Again, I do not find any statistically significant difference in the rate of MIs between patients receiving Celebrex 800 mg/daily and patients receiving non-selective NSAIDs.

65. In addition to my main analysis, I also performed a number of sensitivity analyses using the MIs measure. First, I only conducted this analysis on studies which do not require imputation. [29] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each of the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 13 (Risk Ratio) and Appendix D Exhibit 14 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. There is no statistically significant difference in the rate of MIs for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving non-selective NSAIDs.

iv. Analysis of CV Death

66. Exhibit 13 presents the results for Deaths based on my primary analysis (MH random effects model). For studies including patients receiving Celebrex 200 mg/daily, the point estimate for the risk ratio (RR) is 0.89 (Celebrex over non-selective NSAIDs). The corresponding 95% confidence interval is (0.48, 1.63) which contains the null value one. For risk difference (RD), the point estimate is - 0.0004 and its 95% confidence interval is (-0.0013, 0.0005), which contains the null value zero. Thus, the 95% confidence interval suggests that the difference is not statistically significant leading me to conclude that there is no evidence of any difference in the rate of Deaths between patients receiving Celebrex 200 mg/daily and patients receiving non-selective NSAIDs. For studies including patients receiving Celebrex 400 mg/daily, the point estimate of RR is 0.73 with a 95% confidence interval of (0.35, 1.54), which contains the null value one. For RD, the point estimate is -0.0009 with a 95% confidence interval of (-0.0017, 0.0000), which contains the null value zero. Again, I do not find any statistically significant difference in the rate of Deaths between patients receiving Celebrex 400 mg/daily and patients receiving non-selective NSAIDs. For the 800 mg/daily, the point estimate of RR is 0.78 with a 95% confidence interval of (0.29, 2.09). For RD, the point estimate is -0.0005 with a 95% confidence interval of (-0.0022, 0.0013). Again, I do not find any statistically significant difference in the rate of Deaths between patients receiving Celebrex 800 mg/daily and patients receiving non-selective NSAIDs.

67. In addition to my main analysis, I also performed a number of sensitivity analyses using the Deaths measure. First, I only conducted this analysis on studies which do not require imputation. [30] Second, I also conduct a fixed imputation by adding 0.5 to the zero-arm studies as described in the methodology section. Furthermore, for each the imputation methods (no imputation, fixed imputation, and proportional imputation), I estimated the overall effect using four different estimation methods - MH fixed, MH random, IV fixed, and IV random models. All the results from these analyses are presented in Appendix D Exhibit 15 (Risk Ratio) and Appendix D Exhibit 16 (Risk Difference). I find that the results are very similar to the results based on the MH random model. The sensitivity analysis confirms my conclusions. There is no statistically significant difference in the rate of Deaths for patients receiving Celebrex 200 mg/daily or 400 mg/daily or 800 mg/daily versus patients receiving non-selective NSAIDs.

VI. COMPENSATION

68. I am being compensated at the rate of $450 per hour for my time incurred on this matter. Some of the analyses underlying this report were performed by individuals at Analysis Group, Inc., working under my direction and direct supervision at standard hourly rates charged by that firm. None of this compensation is contingent on the findings or on the outcome of this litigation.

Footnotes

Back to top

Additionally, I have been asked by counsel to review Dr. Muhammad Mamdani's meta-analysis. My review of Dr. Mamdani's meta-analysis is presented in Appendix E.

For descriptions of these methods, see Egger, M., G. Davey Smith, D. Altman, Systematic Reviews in Health Care: Meta-Analysis in Context, British Medical Journal, London (2003) Chapter 15.

The value of 0.5 will be added to those patients receiving drug A who experienced a CV event (raising it from 1 to 1.5); those receiving drug A who did not experience an event (raising it from 49 to 49.5); those receiving drug B who experienced a CV event (raising it from zero to 0.5); and those receiving drug B who did not experience an event (raising it from 50 to 50.5). Since 0.5 "patients" have been added to each group, the total number of "patients" in each drug arm for the purpose of calculating a risk ratio is now 51.

The technique was first developed by Clopper, C., S. Pearson, The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial, Biometrika (1934) Vol. 26, pp. 404-413.

Singh, K., M. Xie, W.E. Strawderman, Combining Information From Independent Sources Through Confidence Distributions, The Annals of Statistics (2005) Vol. 33, No. 1, pp. 159-183.

Right Track contains Pfizer-sponsored and non-Pfizer-sponsored studies. The latter database is called Non-Registry.

Source file is "A319 RTII.xls". Pfizer provides a unique identifier number to its products. A319 corresponds to celecoxib, the generic name for Celebrex.

I eliminated 93 protocols that had been "cancelled," protocols that are at a "concept" or "planned" stage, protocols and studies that are "ongoing," and one study of "unknown legacy." These studies were eliminated because they do not provide all the information that is required to perform my analyses.

For the remaining 126 studies, I examined the Starbase database (Celebrexinventoryfina1_2_24_2005.xls) which has more descriptive titles than Right Track. This step provided information about whether these protocols or studies were double-blind, randomized, and placebo or NSAID controlled. From these 126 reports I was able to identify 43 studies that could potentially be used for my analyses. I plan on reviewing these reports once they are available to me.

I excluded one juvenile rheumatoid arthritis study because my analyses focus on the adult population only.

Nonregistry cox-2 all studies28Jan05.xls

All Status Report - Updated Celebrex Interventional IIRs 23Mar07.xls

44 IIR Protocols.xls, NQ4-00-02-019, and 221_ALS.

<http://www.nih.gov>, visited March 29, 2007.

NIH Website Celecoxib Search Results.pdf

Out of the 73 studies I provided to Dr. Milton Packer, he combined two protocol numbers into one because they corresponded to the same clinical trial (IQ5-97-02-001/EQ5-98-02-002). Additionally, he excluded eight other studies. For three of these studies rofecoxib was the only control group (N49-00-02-181, N49-01-02-145, N49-99-02-149), and for five of these studies there was not enough information in the reports to adjudicate the adverse events (20050305, COXA-0508-261, COXAON-0509-050, 1G20010824, NQ4-00-02-019).

For all the studies (except APC, PreSAP, ADAPT) I use the number of CV events as provided to me by Dr. Packer. For ADAPT, APC, and PreSAP studies, I use the count of adverse events from the publications ADAPT Research Group, Cardiovascular and Cerebrovascular Events in the Randomized, Controlled Alzheimer's Disease Anti-Inflammatory Prevention Trial (ADAPT), PLOS Clinical Trials (2006), and Solomon, S., et al., Effect of Celecoxib on Cardiovascular Events and Blood Pressure in Two Trials for the Prevention of Colorectal Adenomas, Circulation, The Journal of the American Heart Association (2006) Vol. 114,

For example, see Sweeting, M.J., A.J. Sutton, P.C. Lambert, What to Add to Nothing? Use and Avoidance of Continuity Corrections in Meta-Analysis of Sparse Data, Statistics in Medicine (2004) Vol. 23, pp. 1351-1375.

See Greenland S., J. Robins, Estimation of a Common Effect Parameter From Sparse Follow-Up Data, ,Biometrics (1985) Vol. 41, pp. 55-68.

DerSimonian, Rebecca, Nan Laird, Meta-Analysis in Clinical Trials, Controlled Clinical Trials (1986) Vol. 7, pp. 177-188.

By study (or study arms in my exhibits) I am referring to the total number of treatment comparisons, since some studies may compare more than one Celebrex dosages to placebo.

Back to top

22 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

23 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

24 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

25 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

26 By study (or study arms in my exhibits) I am referring to the total number of treatment comparisons, since some studies may compare more than one Celebrex dosages to placebo.

27 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

28 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

29 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

30 For RR, all studies with positive events in both arms do not require imputation. For the RD studies with positive event in at least one arm do not require imputation.

End of Document

Related documents

Selected topics | Secondary Sources | Briefs | Trial Court Documents

Deception Constituting Fraud, and Liability Therefor
Necessary Elements of Fraud Claim